

J Psycholinguist Res (2010) 39:35–49
DOI 10.1007/s10936-009-9124-4

Re-evaluating the Time Course of Gender and Phonological Encoding During Silent Monitoring Tasks Estimated by ERP: Serial or Parallel Processing?

Christian Camen · Stéphanie Morand ·
Marina Laganaro

Published online: 31 July 2009
© Springer Science+Business Media, LLC 2009

Abstract Neurolinguistic and psycholinguistic studies suggest that grammatical (gender) and phonological information are retrieved independently and that gender can be accessed before phonological information. This study investigated the relative time courses of gender and phonological encoding using topographic evoked potentials mapping methods. Event-related brain potentials (ERPs) were recorded using a high resolution electroencephalogram (EEG) system (128 channels) during gender and phoneme monitoring in silent picture naming. Behavioural results showed similar reaction times (RT) between gender and word onset (first phoneme) monitoring, and longer RT when monitoring the second syllable onset. Temporal segmentation analysis (defining dominant map topographies using cluster analysis) revealed no timing difference between gender monitoring and word onset monitoring: both effects fall within the same time window at about 270–290 ms after picture presentation. Monitoring a second syllable onset generated a later effect at about 480 ms. Direct comparison between gender and first phoneme monitoring revealed a difference of only 10 ms between tasks at approximately 200 ms. Taken together, these results suggest that lemma retrieval and phonological encoding may proceed in parallel or overlap. Word onset is retrieved simultaneously with gender, while the longer RT and the later ERP effect for second syllable onset reflect that segmental encoding continues incrementally to the following phonemes.

Keywords Word production · Syntax · Phonology · Time course · ERP · EEG

C. Camen · S. Morand · M. Laganaro
Division of Rehabilitation, University Hospital & University of Geneva, Geneva, Switzerland

S. Morand
Department of Psychology, University of Glasgow, Glasgow, UK

M. Laganaro (✉)
FLSH, ISCL, Logopédie, University of Neuchâtel, Grp NeuroPsychoLinguistique,
1, Espace Louis-Agassiz, 2000 Neuchâtel,
Switzerland
e-mail: marina.laganaro@unine.ch

Introduction

Speaking involves several processes that translate an abstract, conceptual representation into a phonological form. Most models of speech production distinguish three core processes. The first one relates to conceptual-semantic encoding, according to which the targeted conceptual representation is activated. In the second process, the appropriate lexical entry—the lemma—is selected among other co-activated lemmas, including access to words' syntactic properties. The third process deals with the retrieval of the phonological code of the selected word. The retrieved phonological code is then input to a set of phonological and phonetic operations that compute the articulatory motor program.

One crucial question in psycholinguistic research is the time course of these processes, i.e. whether they follow each other without temporal overlap or whether several processes can take place in parallel. A second and related question is the duration of each encoding process. We will focus here on the relative time course of syntactic and phonological encoding of single words.

If all models postulate the existence of several levels of representation, they differ concerning the independence of these levels and the way activation flows between them. For some psycholinguistic (Levelt 1989; Levelt et al. 1999; Schriefers et al. 1990) and neuropsychological models (Ellis et al. 1994), the retrieval of the lemma and the lexeme (phonological information) are independent and are done successively in time. In these discrete models, the phonological form can be encoded only once the semantic and syntactic information have been entirely retrieved. Therefore, these two processes are thought to take place consecutively in time. On the opposite, interactive activation models (Stemberger 1985; Dell 1986, 1988) postulate an interaction between the different encoding processes and levels: lexical and phonological nodes receive activation from all activated semantic nodes and phonological activation feeds back to the semantic level before selection is made. From a temporal perspective, this means that the different encoding processes do not have separate time-courses or that there is a major overlap between these processes.

So, a central point of theoretical disagreement concerns the extent to which lemma selection precedes phonological encoding. There are some psycholinguistic, neurolinguistic and electrophysiological empirical evidences supporting the fact that lexical-syntactic encoding is independent of phonological encoding. The first class of evidence is issued from tip-of-the-tongue (TOT) state experiments. In a TOT state, speakers can report lexical-syntactic properties of a word they are searching for without being able to report the phonological form. Vigliocco et al. (1997), Miozzo and Caramazza (1997a) and Caramazza (1997) showed that Italian speakers were able, in a TOT state, to correctly report the grammatical gender of the words they could not produce, even if they could not provide any correct guess concerning the word form (i.e. number of syllables or phonemes).

Fact is that, as mentioned by Schriefers and Jescheniak (1999) and Caramazza (1997) himself, guesses of grammatical and phonological properties in TOT states are off-line judgments, which do not reflect the availability of information in real time. These results about TOT states only demonstrate that there are separate representations for lexical-syntactic and phonological information, but they do not inform about the temporal organization of lexical processing in language production.

Evidence issued from cognitive neuropsychology and neurolinguistic studies also showed a dissociation between the availability of grammatical and phonological information during lexical access. Henaff-Gonon et al. (1989) were the first to formally observe that subjects with a brain damage suffering from anomia seemed to know the gender of the name they could not produce. Their French speaking patient GM spontaneously produced the correct

gender-marked article of nouns he was unable to produce. [Badecker et al. \(1995\)](#) and [Miozzo and Caramazza \(1997b\)](#) reported a similar observation with a controlled material in terms of the phonology/gender relationship with Italian speaking aphasic patients. Taken together, these neuropsychological observations also indicate that the retrieval of the grammatical feature of a word can proceed independently of the retrieval of its phonological form, but, as already mentioned above for the TOT states, these observations do not allow to determine whether these encoding processes take place successively or in parallel.

The hypothesis of serial encoding of lexical-syntactic information (gender) and phonological information has received support from an electrophysiological study by [Van Turennout et al. \(1998, 1999\)](#). In their experiment, subjects performed a go/no-go task according to two dimensions—syntactic and phonological. The syntactic classification involved the determination of the definite article of the noun (grammatical gender); participants were asked to decide whether the picture corresponded to a *de* or a *het* word (definite determiners for common—masculine and feminine and neuter genders in Dutch). The phonological classification involved the categorization of the initial phoneme; participants were asked to decide whether the name of the picture started with a /b/ or a /s/. The results were based on the analysis of the Lateralized Readiness Potential (LRP), which appears as soon as the primary motor cortex is activated when a decision about response hand has been made, and showed that syntactic information was available 40 ms earlier than phonological information.

Other event-related brain potentials (ERPs) studies using similar paradigms analysed the serial encoding of conceptual/semantic and phonological information ([Schmitt et al. 2000](#); [Rodriguez-Fornells et al. 2002](#); [Jescheniak et al. 2002](#)) or of semantic and syntactic processes ([Schmitt et al. 2001a, b](#)). However, a few ERP studies using similar paradigms supported the opposite view, i.e. parallel encoding processes during speech production ([Abdel-Rahman and Sommer 2003](#); [Abdel-Rahman et al. 2003](#)). [Abdel-Rahman and Sommer \(2003\)](#) analysed the time course of semantic and phonological processes with the go/no-go paradigm and manipulated the duration of semantic processing by increasing the classification difficulty (real size or diet), in order to test whether the duration of semantic processing affected the beginning of phonological encoding, as predicted by serial models. They analysed the response to the LRP interval (R-LRP) as an estimate of the duration of semantic and phonological encoding. The results supported the parallel encoding since the interval between the semantic and the phonological task decreased when difficulty increased instead of the opposite.

In the context of tracking the time course of these processes, another point must be addressed concerning the go/no-go paradigm studies. The two choice go/no-go paradigm is well suited to determine the serial or parallel nature of the processes but doesn't allow a direct access to the time course of these processes for the following reasons. Firstly, the latency of the LRP is taken per se and reveals the moment at which the information, once available, is transmitted to pre-motor cortex, but not the moment at which the information is processed ([Van Turennout et al. 1997, 1998](#); [Schmitt et al. 2000, 2001a, b](#); [Rodriguez-Fornells et al. 2002](#)). Secondly, it involves a computation (subtraction) in order to obtain the relative timing of syntactic, semantic, or phonological processes ([Abdel-Rahman and Sommer 2003](#); [Abdel-Rahman et al. 2003](#)); the authors calculate the difference between the response time and the moment at which the syntactic or phonological information is available (cued by the latency of the LRP, R-LRP). But once again this computation reveals the moment at which the syntactic or phonological information is available, not the moment at which the processing of this kind of information takes place. So the results of these studies are only informative about the termination of these processes and do not provide direct timing information about the processes themselves.

In the present experimental design we tried to directly address the moment at which these two types of information are processed. As already done in the mentioned EEG studies on language production, we will investigate single word production with a silent monitoring task. Monitoring tasks have been largely used to investigate language production processes, especially at the level of phonological encoding (Wheeldon and Levelt 1995; Wheeldon and Morgan 2002; Schiller et al. 2003, 2006). The issue of whether monitoring tasks are performed on the generated production code or on a subarticulated production via a perceptual monitoring mechanism has been addressed by Wheeldon and Levelt (1995), who asked participants to monitor for a specific segment while performing an articulatory suppression task. Their results showed that the addition of an articulatory suppression task had only a limited effect on phoneme monitoring. Since the articulatory suppression task is thought to prevent phonetic encoding, these results suggested that monitoring can be done on a representation that is not sub-articulated, that is, on an abstract code. Other evidence showing that speakers can monitor an abstract representation without having recourse to sub-articulation comes from the fact that articulatory suppression only had a limited effect in silent rhyme and homophony judgment tasks (Besner 1987) and from the observation that phoneme monitoring slows down at syllable boundaries, which seems to reflect an online production process (Jansma and Schiller 2004).

Method

Task

We used a simple monitoring paradigm in which subjects separately monitored for syntactic/grammatical and phonological information on the same material. In one condition subjects had to silently monitor if the name corresponding to the picture was congruent with the proposed gender. In a second condition, subjects monitored if the name corresponding to the picture contained the target phoneme. Both conditions required positive and negative responses, contrary to a go/no-go paradigm in which no negative responses are given. This paradigm, and the manner with which the EEG signal is then processed, allowed to determine the moment at which a difference occurred between yes (go) and no response in each condition. By comparing the conditions, we should be able to estimate (1) the time-window of the intra-condition yes/no differences, (2) the time-windows of inter-condition yes (go) responses, in order to determine if the processes are done in parallel or serially.

Participants

Fourteen subjects participated in the experiment. All were healthy French native speakers, right handed (8 women, 6 men), aged 19–34 (mean \pm S.D. = 27 ± 6.6) and were paid for their participation. They were informed about the ERP procedure study and gave a written consent.

Materials

A set of 40 words (20 bi-syllabic and 20 tri-syllabic) and their corresponding black and white drawings were selected in the Alario et al. (2004) and Bonin et al. (2003) databases. All the items had name agreement superior to 65% and were selected according to the onset of their first or second syllable, which led to a total of 11 target phonemes occurring an equal number of times in each position (first or second syllable onset).

Procedure

Each participant underwent two experimental conditions in a counterbalanced order, a gender monitoring and a phoneme monitoring task. Before starting the experiment and before receiving any further instructions, the participants were familiarized with the material: they were given a booklet including all drawings and their names. They were asked to examine all the drawings and to check whether the name they would use corresponded to the proposed name.

Subjects sat 0.7 m in front of the screen, viewing pictures of approximately 0.09 m with a visual angle of 3.67°.

Gender

In the ‘gender’ monitoring task participants had to silently monitor for the grammatical gender of the word corresponding to the picture. French distinguishes two classes of grammatical gender, masculine and feminine nouns (other languages might have an additional class—neuter gender—, like German, or no grammatical gender, like English). It is important to note that grammatical gender does not have an unequivocal link to morpho-phonological forms in French (for example, “cigare” (cigar) is masculine, the determiner is “le”, but “guitare” (guitar) is feminine, the determiner is “la”). Auditory instructions were given at the beginning of each experimental block (“*on dit LE...*”, “*one says the*”—*LE is the French article for masculine nouns*), followed by presentation of a block of about 11 items (with 2–3 fillers at the beginning of each block to ensure that subjects correctly understood the new instruction), with a total of 10 alternated blocks. For each of them, subjects had to respond on a button box, with index for positive responses and major for negative responses. After each block a new auditory instruction indicated the gender (“*on dit LA...*”—*one says the* —*LA is the French article for feminine gender*). A trial in a block began with a cross in the middle of the screen for 500 ms, then the picture appeared on the screen during 2,500 ms. Response latencies (RT) were recorded from the moment the picture appeared on the screen. Each item was presented 2 times, one in a positive and one in a negative condition, for a total of 80 analyzed trials (40 positive and 40 negative trials, without counting the 30 fillers) per subject.

Phoneme Monitoring

Participants monitored for a target phoneme in a silent picture-naming task. The procedure followed closely the one described for gender monitoring. The target phoneme was given auditorily before each block and, to avoid any confusion, it was always exemplified into a first name beginning with the target consonant (e.g. “[*k*] like Kim”). Participants had to monitor for the presence of the target phoneme in any position in the word within blocks of about 13 trials (with 2–3 fillers at the beginning), with a total of 15 alternated blocks (which led to a total of 195 trials including fillers). The target phoneme and the orthographic form of the words to be monitored did not always correspond, for example, subjects monitored for the phoneme /s/ in “serpent” (snake), but also in “citron” (lemon) and “pin^{ce}au” (brush), or for /f/ in “fourmi” (ant), and in “dauphin” (dolphin). 15 fillers were also selected from the same databases. In the positive condition, the target phoneme was located either at the onset of the first syllable (*pho1*) or at the onset of the second syllable (*pho2*) of the target word. Each item was presented 4 times, two in the positive (*pho1* and *pho2*, different phonemes) and two in the negative condition, for a total of 160 analyzed trials (40 *pho1* positive trials,

40 pho2 positive trials and 80 negative trials, without counting the 35 fillers) per subject. Stimuli were presented in a pseudo-random order in order to guarantee an even distribution of conditions (yes-no response, first-second syllable onset,...) across the blocks of trials.

Electrophysiological Recordings and Data Analysis

The electroencephalogram (EEG) was recorded continuously using the Active—Two Biosemi EEG system (Biosemi V.O.F Amsterdam, Netherlands) with 128 channels covering the entire scalp. Signals were sampled at 512 Hz in a bandwidth filter of 0–314 Hz. EEG data were band-pass-filtered to 1–30 Hz. Correct response trials were visually inspected, and trials contaminated by eye-blinks and/or artifacts were rejected and excluded from averaging. Artifacted EEG channels were interpolated. For each task, positive and negative trials were averaged separately.

For all analyses, baseline was defined as the 100 ms pre-stimulus period. RT superior to 2,000 ms and wrong responses were excluded from the analyses. All the participants performed the tasks very accurately; consequently no subject was excluded due to behavioural performance. Subjects who did not reach the criteria of a minimum of 30 averaged epochs per condition after rejection of artefacts and eye-blinks were excluded and replaced.

The ERPs were submitted to analyses of the electric field at the scalp. A topographic (map) pattern analysis (segmentation) was computed on compared datasets. Maps were compared over time within and between conditions, since topographic changes indicate differences in the brain's active generators (Fender 1987). This method is independent of the reference electrode (see e.g. Michel et al. 2001, 2004) and it is insensitive to pure amplitude modulations across conditions (topographies of normalized maps are compared). A modified cross-validation criterion determined the number of maps that explained the whole group-averaged data set (Pascual-Marqui et al. 1995; Michel et al. 2001). Statistical smoothing was used to eliminate temporally isolated maps with low strength. Additionally, a given topography had to be present for at least 10 time frames (20 ms). This procedure is described in detail in Pascual-Marqui et al. (1995, see also Murray et al. 2008). This approach allows to summarize ERP data into a limited number of electrocortical map configurations and identifying time periods during which different conditions evoke different electrocortical configurations.

The patterns of maps observed in the group-averaged data was statistically tested by comparing each of these maps with the moment-by-moment scalp topography of individual subjects' ERPs from each condition. Each time point was labelled according to the map with which it best correlated spatially, yielding a measure of map presence that was in turn submitted to an ANOVA with condition (positive/negative) and maps (hereafter referred to as 'fitting'). This fitting procedure revealed whether a given experimental condition was more often described by one map versus another, and therefore if different generator configurations better accounted for particular conditions (i.e. if there is a significant interaction between maps and conditions). This approach has been validated in other cognitive domains (Ducommun et al. 2002; Murray et al. 2006; Pegna et al. 2002; Schneider et al. 2007). These analyses were performed using the Cartool software (<http://brainmapping.unige.ch/Cartool.php>).

Results

Behavioural Results

Mean reaction times (RT) and accuracy in each task are shown in Table 1. RT comparison between 'gender' and 'phoneme' conditions showed no difference between gender

Table 1 Mean reaction time (RT, in milliseconds) and percent correct for gender and phoneme monitoring tasks (positive conditions) and standard deviations (in brackets)

	Gender	Pho1 (first syllable onset)	Pho2 (second syllable onset)
RT	883 ms (134)	849 ms (111)	1,003 ms (126)
Accuracy	96.4% (2.9%)	96.8% (2.6%)	91.4% (5.6%)

monitoring and monitoring for a phoneme at word onset position (paired t-tests calculated by participants (t_1) and by items (t_2): $t_1(13) = 1.28$, $p = 0.22$, $t_2(39) = 1.92$, $p = 0.063$) and a significant difference of 120 ms between gender and second syllable onset ($t_1(13) = 4.39$, $p < .001$, $t_2(39) = 6.37$, $p < .001$).

Electrophysiological Results

The topographic pattern analysis identified periods of stable electric field configurations at the scalp and determined whether different configurations of brain generators accounted for responses to ‘phoneme’ and ‘gender’ monitoring. In the intra-task analysis, the comparison was carried out between the positive and negative responses. In the inter-task analysis, comparison was carried out between gender and first phoneme monitoring positive trials. From 6 to 10 different topographies accounted for the data set during the 600 ms post stimulus period.

Intra-task Analyses

Gender Six different topographies accounted for the whole data set (i.e., the 600 ms post stimulus period from both experimental conditions) with a global explained variance of 90.7%. Comparing the positive and negative responses of the ‘gender’ condition revealed identical topographies with different time distribution over the 270–600 ms period (Fig. 1a). The fitting procedure statistically confirmed this differential time distribution, yielding an interaction for duration between conditions and maps 4 and 5 ($F(1, 13) = 5.17$, $p < .05$), with neither main effect reaching a 0.05 significance criterion. Direct comparisons between conditions on duration were significant for each map (respectively, map 4: $t(13) = 2.27$, $p < .05$; map 5: $t(13) = 2.27$, $p < .05$). Analyses of Global Explained Variance (GEV) revealed no significant interaction between conditions and maps 4 and 5 ($F(1, 13) = 1.2$, $p = .28$).

The fitting procedure yielded a second interaction on duration between conditions and maps 5 and 6 ($F(1, 13) = 7.3$, $p < .05$), with neither main effect. Direct comparisons of duration between conditions on duration were significant for each map (respectively, map5: $t(13) = 2.7$, $p < .05$; map 6: $t(13) = 2.7$, $p < .05$). In terms of GEV, the fitting procedure showed no significant interaction ($F(1, 13) = 1.6$, $p = .22$). Once again maps were differently distributed over time, with map 5 slightly longer in negative than positive responses.

In sum, over the 270–600 ms period, identical maps were active but differently distributed over time, with a first difference between positive and negative responses occurring between 270 and 290 ms.

Phoneme Ten different topographies accounted for the whole data set with a global explained variance of 93.73% (over the 600 ms post stimulus period from the three experimental conditions; *pho1*, *pho2* and negative responses).

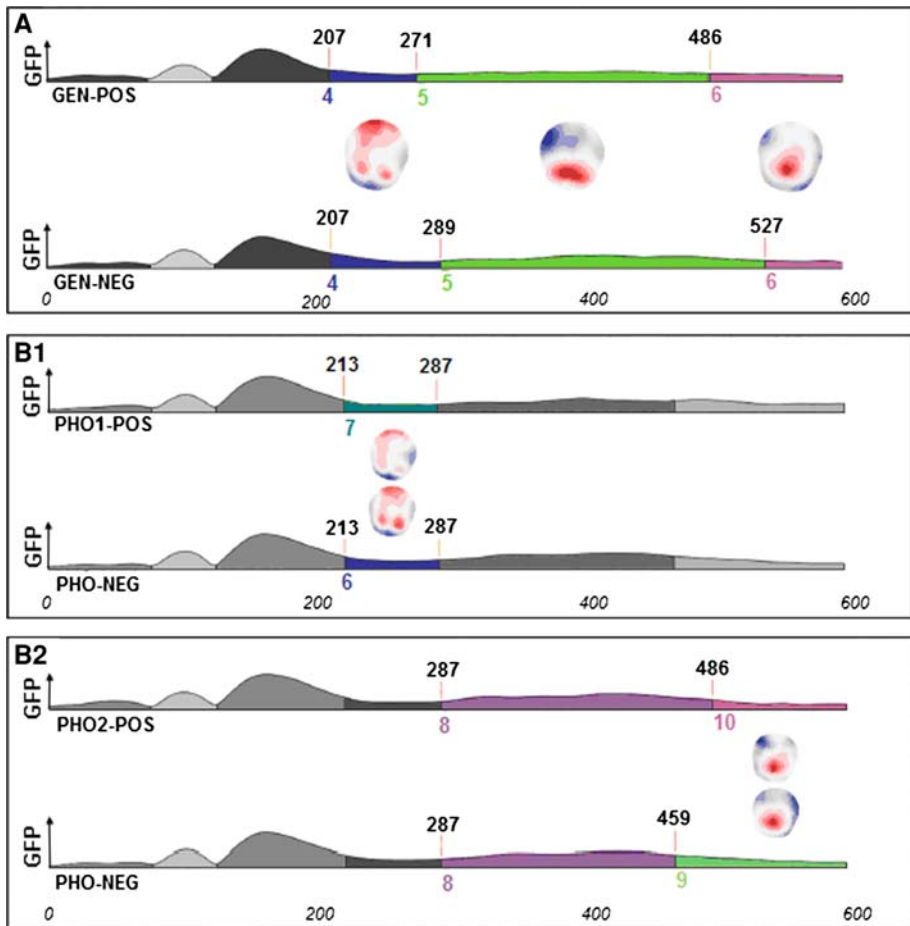


Fig. 1 Segmentation results for the contrasts between conditions from 0 to 600 ms (X-axis). Y-axis represents Global Field Power (GFP). Time periods and topographic maps are indicated for significant differences. **a** Gender monitoring, positive versus negative. **b1** Phoneme 1 (word onset) monitoring, positive versus negative. **b2** Phoneme 2 (second syllable onset) monitoring, positive versus negative

Two statistical comparisons were conducted on the basis of this triple segmentation, concerning the phoneme located on the onset of the first syllable (*pho1*) versus negative responses and the phoneme located on the onset of the second syllable (*pho2*) versus negative responses.

- 1) Across positive responses for ‘pho1’ and negative responses, different topographies were identified over the 213–287 ms period (Fig. 1b1). A substitution of maps occurred over this 213–287 ms period. The fitting procedure confirmed this observation, yielding an interaction between conditions and maps 6 and 7 in terms of duration over this time period ($F(1, 13) = 18.36, p < .001$), with neither main effect reaching our 0.05 significance criterion. Direct comparisons of duration between conditions were significant for each map (respectively, map 6: $t(13) = 4.28, p < .001$; map 7: $t(13) = 4.28, p < .001$). Analyses of GEV revealed a significant interaction between conditions and these maps ($F(1, 13) = 8.02, p < .05$). Direct comparisons of GEV between conditions were

significant for each map (respectively, map 6: $t(13) = 3$, $p < .05$; map 7: $t(13) = 3$, $p < .05$), confirming the observation of a substitution of maps, with map 6 dedicated to negative and map 7 to positive responses.

- 2) Across positive responses for *pho2* and negative responses, two topographies were identified over the 486–600 ms period (Fig. 1b2). Map 8 was differently distributed over the 287–486 ms period, as indicated by a significant direct comparison of duration ($t(13) = 2.66$, $p < .05$). Analyses of GEV on this contrast revealed no significant difference ($t(13) = 1.03$, $p = .32$). The fitting procedure revealed a significant interaction between conditions and maps 9 and 10 in terms of duration ($F(1, 13) = 18.07$, $p < .001$). Direct comparisons between conditions on duration were significant for each map (respectively, map 9: $t(13) = 4.25$, $p < .001$; map 10: $t(13) = 4.25$, $p < .001$). Analyses of GEV revealed a significant interaction between conditions and these maps ($F(1, 13) = 14.74$, $p < .01$). Direct comparisons of GEV between conditions on GEV were also significant for each map (respectively, map 9: $t(13) = 3.11$, $p < .01$; map 10: $t(13) = 3.32$, $p < .05$), confirming that different maps predominated in the positive and negative responses, with map 9 dedicated to negative and map 10 to positive responses.

In sum, these comparisons revealed that different maps predominated in the positive and negative responses with a substitution between 213 and 287 ms for positive *pho1* versus negative responses and at 486 ms for positive *pho2* versus negative responses.

Inter-task Analysis: Gender Versus Phoneme Monitoring

A direct comparison between positive responses at first phoneme and gender monitoring was carried out since this comparison is crucial for the question of the relative time course of gender and phonological encoding. Seven different topographies accounted for the whole data set with a global explained variance of 94.38%. Comparing the positive trials of *pho1* and gender revealed identical topographies with differential time distribution over the 123–266 ms period (Fig. 2). The fitting procedure confirmed this observation, yielding an interaction in terms of duration between conditions and maps over this 123–266 ms period ($F(1, 13) = 7.9$, $p < .05$). Direct comparisons of duration between conditions were significant for each map (respectively, map 1: $t(13) = 2.81$, $p < .05$; map 4: $t(13) = 2.81$, $p < .05$), with map 1 shorter and map 4 longer in the gender than in the phoneme condition. In terms of GEV, the fitting procedure showed no significant difference. The interaction and contrasts in terms of duration provided information about the fact that these two processes shared the same topographies from 123 to 266 ms but were differently distributed over time.

In sum, a direct comparison between positive trials of *pho1* and gender revealed identical topographies with differential time distribution over the 123–266 ms period. These analyses revealed a difference of only 10 ms between tasks at approximately 200 ms.

Discussion

We used high temporal resolution evoked potentials and a segmentation procedure to estimate the time course of syntactic and phonological encoding processes during silent gender and phoneme monitoring tasks. Both grammatical gender and initial phoneme encoding yes/no effects arose during the same time window at about 270–290 ms, while direct comparison between gender and phoneme monitoring on yes responses revealed similar maps, with a 12 ms earlier onset at about 200 ms for gender than for first phoneme.

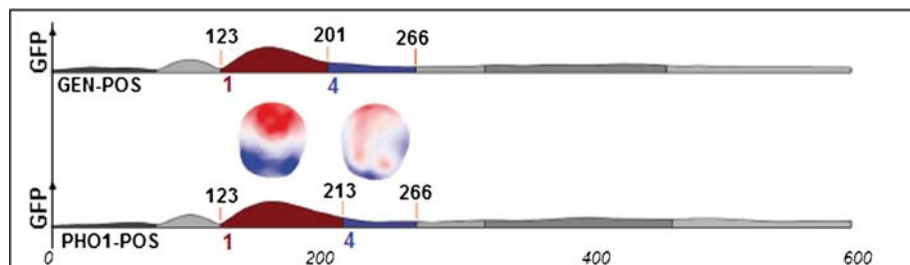


Fig. 2 Segmentation results for the contrasts between gender (GEN) and first phoneme (PHO1) monitoring (positive trials only) from 0 to 600 ms (X-axis). Y-axis represents Global Field Power (GFP). Time periods and topographic maps are indicated for map1 and map 4

Differently from previous studies, we carried out intra-task comparison between positive and negative responses and inter-task comparison between positive responses (corresponding to go-responses in the traditional paradigms) with spatio-temporal segmentation analyses. The different comparisons associated with the spatio-temporal segmentation procedure directly tap into the time windows in which presence of the targeted information was detected by the subjects, avoiding the uncertainty linked to motor response/preparation of the LRP analyses. This allows to show, across all electrodes sites, the exact time window during which brain activity differs. The identified periods of stable topographies are thought to reflect functional microstates linked to specific cognitive processes.

Firstly, the time window yielding a difference between positive and negative responses was quite similar for the grammatical gender and the first phoneme monitoring, both falling around 270–290 ms, although the nature of the differences varied between these two comparisons. In the gender task, the same topographies with different time distribution were observed in the positive and in the negative condition, whereas a map substitution between positive and negative conditions arose in phoneme monitoring task. The nature of the task itself—determining the grammatical gender or determining the presence of a target phoneme—could be involved in this difference. Indeed, even if the task was a two-choice decision in both cases, gender basically has only two values in French (masculine or feminine), while detecting the presence of a target phoneme in whichever position has potentially more possible issues. Secondly, direct comparison between gender and first phoneme confirmed these results and did not reveal a clear serial pattern of maps. Sequential encoding of information would predict that, at least, one map in one condition entirely precedes the corresponding map in the other condition, which was not the case in our results. The spatio-temporal segmentation showed that the same topographic maps, i.e. the same network of intracranial generators, appeared in the positive trials of the two tasks, with only a shift of 10 ms. Finally, there was no RT difference between positive responses for first phoneme and gender monitoring. At first sight the observed gender monitoring RTs (mean of 883 ms) seem quite long, but they are very similar to those reported by Navarrete et al. (2006, exp. 2, mean RT around 900 ms) with a silent gender decision task, where subjects had to press one of two buttons for pictures corresponding to masculine or feminine words. Moreover, the monitoring for a phoneme located on the second syllable onset lead to significantly longer RT, indicating that gender and first phoneme monitoring do not need access to the phonological information of the second syllable.

Hence our behavioural and electrophysiological results suggest that processing of first phoneme and grammatical gender monitoring do not occur in a complete serial way; these

two processes seem to take place in parallel or to overlap and gender monitoring may start only 10 ms before phoneme monitoring.

These results are clearly at odds with other electrophysiological results supporting serial processing, like in the [Van Turennout et al. \(1998\)](#) study, who found a 40 ms difference in LRP between syntactic and phonological encoding. But, as suggested in the Introduction, a methodological point must be addressed concerning these go/no-go studies. In [Van Turennout et al. \(1998\)](#), pictures were classified as common gender “de” or neuter “het” by a key press with the left (common) or right (neuter) hand which, however, had to be executed in function of word onset (if the object’s name began, for example, with a /b/, while the response had to be withheld (no-go condition) when the first sound was a /s/). An LRP was expected in go and in no-go condition because, for this latter condition, the information about the non-execution of the response (in this design the phonological information) was available, according to the theory, later than decision about syntactic information. When decision was linked to phonological information and execution linked to syntactic information (reversed design), no LRP was expected in the no-go condition because syntactic information was available earlier than phonological information. In this context, no preparation is activated, so no LRP is expected. The authors claimed that these LRP findings ([Van Turennout et al. 1999](#)) show that syntactic information is available earlier than phonological information, confirming serial or cascaded models of processing in language production, in contrast to parallel processing. However, an alternative explanation also holds for these results. Indeed, as stated by [Abdel-Rahman and Sommer \(2003\)](#), if syntactic retrieval starts simultaneously with phonological encoding but terminates earlier, the same observations in LRPs, as described above, are predicted. Therefore the 40 ms difference observed in LRP between syntactic and phonological encoding is also compatible with a parallel/independent retrieval of syntactic and phonological knowledge with the gender retrieval process lasting shorter than phonological encoding process.

[Abdel-Rahman and Sommer \(2003\)](#) tested the hypothesis of simultaneity by using a similar two-choice go/no-go paradigm to compare semantic versus phonological processing. As mentioned in the Introduction, they additionally manipulated the duration of semantic processing. LRP effect results were clearly in accordance with parallel models. In a second experiment the assignment of semantic and phonological properties to hand and go/no-go decision was reversed. Once again the results confirmed the predictions from parallel models. And finally, in order to ensure that the observed results were not linked to the choice of the semantic dimension and not due to the presence of two separate semantic classification (easy vs. difficult), the same design was used in a subsequent experiment ([Abdel-Rahman et al. 2003](#)), replacing the size or diet by the animacy (a more basic semantic feature) and confirming the parallel predictions.

Although these experiments relate to semantic instead of syntactic processing, they showed, at least, that simultaneous initiation of different encoding is a realistic possibility. Considered jointly with the data presented in the Introduction, it seems plausible that two (or more) encoding processes may take place simultaneously.

Secondly, the timing of the present results is in line with the estimation by [Indefrey and Levelt \(2004\)](#) but apparently earlier than in other ERP studies. According to [Indefrey and Levelt \(2004\)](#), lemma selection should take place between 150 and 250 ms after picture onset and phonological encoding was estimated to begin at about 250 ms. Our results showed that access to gender should be completed in a time window prior to 270 ms. The timing of phonological encoding does also converge with this previous estimation, since the word onset (first phoneme) effect fall between 270 and 290 ms. The time-course revealed with gender and phoneme monitoring is also in line with the time-window of lexical frequency

effect observed with a naming task and similar ERP analyses (Laganaro et al. 2009). In that study waveforms analysis and temporal segmentation indicated an effect of lexical frequency in the time-window between 260 and 330 ms, also suggesting that that lexical selection is accomplished during this time-window. However, our results showed an earlier effect of gender monitoring than other ERP studies. In Van Turennout et al. (1998) gender monitoring generated an effect at about 370 ms and initial phoneme at about 400 ms after picture onset. But, as already mentioned, the development of an LRP does not indicate precisely when information becomes available; it rather indicates that information is used to activate the (motor-) responses. This implies that relevant information is available at least at LRP onset but that it might have been available earlier as well. This means that the 100 ms delay in time observed between the studies could be linked to this timing uncertainty. In our study we probably accessed directly to the time window during which the encoding processes occurred, whereas studies that make use of LRP information allow to observe a motor-related brain activity preceding the overt response. The same reasoning applies to the 464/540 ms LRP effects and to the 550 ms N200 effect observed in Schmitt et al. (2001a, b) studies for semantic and syntactic information. In Schmitt et al. (2001a), the N200 effect (“no-go minus go ERPs”) reflects the upper time limit by which information about whether an actual response needs to be made or withheld must have become available. This time is typically captured by the mean peak latency of the N200 effect. They found that the mean peak latency of the N200 effect was 550 ms when the go/no-go response was contingent on syntactic information, in their case gender information. In Schmitt et al. (2001b), the mean peak latency of the N200 effect was 464 ms when the go/no-go decision was based on gender information. Compared to our superior limit of 270 ms these latencies are relatively late. It has been suggested that the N200 amplitude is a function of neuronal activity required for response inhibition. One interpretation of the N200 to no-go stimuli, receiving empirical support, is that it reflects an inhibitory process emanating from structures in the prefrontal cortex; response inhibition can be viewed as part of the executive control which involves working memory processes mediated by prefrontal cortex. In this sense it is obvious that these prefrontal structures, as pre-motor structures involved in the LRP effect, need a period of time to be activated and that some care has to be taken in interpreting the initial development of an LRP/N200 effect as an indication of the moment in time at which information becomes available.

Monitoring a phoneme located on the onset of the second syllable showed a longer RT than monitoring phoneme located on the first syllable. Moreover, monitoring *pho1* showed a yes/no effect between 213 and 287 ms whereas monitoring *pho2* generated a later yes/no effect at about 460 ms. These results are consistent with previous studies showing longer RT for monitoring phoneme on the onset of the second syllable than for monitoring the onset of the first syllable and suggesting an incrementally encoding or phonological scanning of the segmental information (Wheeldon and Levelt 1995; Wheeldon and Morgan 2002).

In all cases the results were in line with an incremental/serial processing within phonological encoding. According to Indefrey and Levelt (2004) the estimated range for the availability of the first phonological segment is in a time window between 230 and 430 ms and phoneme-by-phoneme self-monitoring proceeds at a speed of about 25 ms per phoneme. Our results fall within this time window, but the timing difference between first phoneme and second syllable onset monitoring is greater than in other studies. Van Turennout et al. (1997, 1998) observed approximately 100 ms time difference between encoding of initial and final phoneme, as cued by the LRP. In our study, the ERP time difference observed between encodings was approximately 250 ms and the mean RT difference was 154 ms. This 150 ms delay between studies could be due to the paradigm we used and the nature of the words used in these ERP studies. First, in our experiment the participant did not know where the target

phonemes (first or second syllable) were, contrary to Van Turenhout et al. (1997, 1998, 1999) studies, in which the position was given (initial or final phoneme for example). In this sense subjects did not need to scan all the phonemes of the word. Moreover, the length of the word stimuli varied between these studies. For example, Van Turenhout et al. (1997) used mostly mono- syllabic and some bi-syllabic words whereas stimuli were bi- and tri-syllabic words in the present study. Jointly considered, and according to the hypothesis of an incremental scanning mechanism involved in monitoring target phonemes, it is plausible that the use of mono-syllabic words explains the shorter difference between initial and final phonemes in their study, especially when target location is known.

In summary, word onset and grammatical gender seem to be retrieved simultaneously, with only a 10 ms earlier onset in gender than in first phoneme monitoring, while longer RT and later yes/no effects for the second syllable onset reflect that segmental encoding continues incrementally to the following phonemes. The very limited timing difference and the lack of a serial pattern between gender and first phoneme suggest that gender retrieval and phonological encoding proceed in parallel or overlap, with only a few ms shift at the beginning of the processes. The experimental procedure described in this article provides a new EEG technique for estimating the time course of the encoding processes implied in speech production, supplementing the LRP approach.

Acknowledgments This research was supported by Swiss National Science Foundation grant no. 105312-108284 and PP001-118969. The Cartool software has been programmed by Denis Brunet, from the *Functional Brain Mapping Laboratory*, Geneva, Switzerland, and is supported by the *Center for Biomedical Imaging (CIBM)* of Geneva and Lausanne (<http://brainmapping.unige.ch/Cartool.php>).

References

- Abdel-Rahman, R. A., & Sommer, W. (2003). Does phonological encoding in speech production always follow the retrieval of semantic knowledge? Electrophysiological evidence for parallel processing. *Cognitive Brain Research*, 16, 372–382.
- Abdel-Rahman, R. A., Van Turenhout, M., & Levelt, W. (2003). Phonological encoding is not contingent on semantic feature retrieval: An electro-physiological study on object naming. *Journal of Experimental Psychology*, 29(5), 850–860.
- Alario, X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, 36, 140–155.
- Badecker, W., Miozzo, M., & Zanuttini, R. (1995). The two-stage model of lexical retrieval: Evidence from a case of anomia with selective preservation of grammatical gender. *Cognition*, 57, 193–216.
- Besner, D. (1987). Phonology, lexical access in reading, and articulatory suppression: A critical review. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 39A, 467–478.
- Bonin, P., Peerman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, & Computers*, 35, 158–167.
- Caramazza, A. (1997). How many levels of processing are there in lexical access?. *Cognitive Neuropsychology*, 14, 177–208.
- Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the 'tip-of-the-tongue' phenomenon. *Cognition*, 64, 309–343.
- Dell, G. S. (1988). The retrieval of phonological forms in production: Test of predictions from a connectionist model. *Journal of Memory and Language*, 27, 124–142.
- Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.
- Ducommun, C. Y., Murray, M. M., Thut, G., Bellmann, A., Viaud-Delmon, I., Clarke, S., & Michel, C. M. (2002). Segregated processing of auditory motion and auditory location: An ERP mapping study. *NeuroImage*, 16, 76–88.

- Ellis, A. W., Franklin, S., & Crerar, A. (1994). Cognitive neuropsychology and the remediation of disorders of spoken language. In M. J. Riddoch & G. W. Humphreys (Eds.), *Cognitive neuropsychology and cognitive rehabilitation*. Hove: Lawrence Erlbaum Associates.
- Fender, D. H. (1987). Methods of analysis of brain electrical and magnetic signals. In A. S. Gevins & A. Remond (Eds.), *Handbook of electroencephalography and clinical neurophysiology*, (Vol.1), (pp. 355–399). Amsterdam: Elsevier.
- Henaff-Gonon, M. A., Bruckert, R., & Michel, F. (1989). Lexicalization in an anomic patient. *Neuropsychologia*, 27, 391–407.
- Indefrey, P., & Levelt, W. (2004). The spatial and temporal signatures of word production components. *Cognition*, 92, 101–144.
- Jansma, B. M., & Schiller, N. O. (2004). Monitoring syllable boundaries during speech production. *Brain and Language*, 90(1–3), 311–317.
- Jescheniak, J. D., Schriefers, H., Garrett, M. F., & Friederici, A. D. (2002). Exploring the activation of semantic and phonological codes during speech planning with event-related brain potentials. *Journal of Cognitive Neuroscience*, 14(6), 951–964.
- Laganaro, M., Morand, S., Schwitler, V., Zimmermann, C., Camen, C., & Schnider, A. (2009). Electrophysiological correlates of different anomic patterns in comparison with normal word production. *Cortex*, 45, 697–707.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–38.
- Michel, C. M., Thut, G., Morand, S., Khateb, A., Pegna, A. J., & Gravede de Peralta, R. (2001). Electric source imaging of human brain functions. *Brain Research. Brain Research Reviews*, 36, 108–118.
- Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Gravede de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology*, 115, 2195–2222.
- Miozzo, M., & Caramazza, A. (1997). The retrieval of lexical-syntactic features in tip-of-the-tongue states. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(6), 1410–1423.
- Miozzo, M., & Caramazza, A. (1997). On knowing the auxiliary of a verb that cannot be named: Evidence for the independence of grammatical and phonological aspects of lexical knowledge. *Journal of Cognitive Neuroscience*, 9, 160–166.
- Murray, M. M., Camen, C., Gonzalez Andino, S. L., Bovet, P., & Clarke, S. (2006). Rapid Brain Discrimination of Sounds of Objects. *Journal of Neuroscience*, 26(4), 1293–1302.
- Murray, M. M., Brunet, D., & Michel, C. (2008). Topographic ERP analyses: A step-by-step tutorial review. *Brain Topography*, 20, 249–269.
- Navarrete, E., Basagni, B., Alario, F. X., & Costa, A. (2006). Does word frequency affect lexical selection in speech production?. *Quarterly Journal of Experimental Psychology*, 59(10), 1681–1690.
- Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1995). Segmentation of brain electrical activity into microstates: Model estimation and validation. *IEEE Transactions on Biomedical Engineering*, 42, 658–665.
- Pegna, A. J., Khateb, A., Murray, M. M., Landis, T., & Michel, C. M. (2002). Neural processing of illusory and real contours revealed by high-density ERP mapping. *NeuroReport*, 13, 965–968.
- Rodriguez-Fornells, A., Schmitt, B. M., Kutas, M., & Münte, T. F. (2002). Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming. *Neuropsychologia*, 40, 778–787.
- Schiller, N. O., Bles, M., & Jansma, B. M. (2003). Tracking the time course of phonological encoding in speech production: An event-related brain potential study. *Cognitive Brain Research*, 17(3), 819–831.
- Schiller, N. O., Jansma, B. M., Peters, J., & Levelt, W. J. (2006). Monitoring metrical stress in polysyllabic words. *Language and Cognitive Processes*, 21, 112–140.
- Schmitt, B. M., Münte, T. F., & Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology*, 37, 473–484.
- Schmitt, B. M., Rodríguez-Fornells, A., Kutas, M., & Münte, T. F. (2001). Electrophysiological estimates of semantic and syntactic information access during tacit picture naming and listening to words. *Neuroscience Research*, 41, 293–298.
- Schmitt, B. M., Schiltz, K., Zaake, W., Kutas, M., & Münte, T. F. (2001). An electrophysiological analysis of the time course of conceptual and syntactic encoding during tacit picture naming. *Journal of Cognitive Neuroscience*, 13, 510–522.
- Schnider, A., Mohr, C., Morand, S., & Michel, C. M. (2007). Early cortical response to behaviorally relevant absence of anticipated outcomes: A human event-related potential study. *Neuroimage*, 35(3), 1348–1355.
- Schriefers, H., & Jescheniak, J. D. (1999). Representation and processing of grammatical gender in language production: A review. *Journal of Psycholinguistic Research*, 28(6), 575–600.

- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86–102.
- Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1). Hillsdale, NJ: Erlbaum.
- Van Turenout, M., Hagoort, P., & Brown, C. M. (1997). Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 787–806.
- Van Turenout, M., Hagoort, P., & Brown, C. M. (1998). Brain activity during speaking: From syntax to phonology in 40 milliseconds. *Science*, 280, 572–754.
- Van Turenout, M., Hagoort, P., & Brown, C. (1999). The time course of grammatical and phonological processing during speaking: Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, 28(6), 649–676.
- Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, 8, 314–317.
- Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language*, 34, 311–334.
- Wheeldon, L. R., & Morgan, J. L. (2002). Phoneme monitoring in internal and external speech. *Language and Cognitive Processes*, 17(5), 503–535.